

Theory and Application of Shapelets to the Analysis of Surface Self-assembly Imaging

Robert Suderman^a, Daniel Lizotte^a, Nasser Mohieddin Abukhdeir^{b,c,*}

^aDepartment of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

^bDepartment of Chemical Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

^cWaterloo Institute for Nanotechnology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

Abstract

A method for quantitative analysis of local pattern strength and defects in surface self-assembly imaging is presented and applied to images of stripe and hexagonal ordered domains. The presented method uses “shapelet” functions which were originally developed for quantitative analysis of images of galaxies ($\propto 10^{20}$ m). In this work, they are used instead to quantify the presence of translational order in surface self-assembled films ($\propto 10^{-9}$ m) through reformulation into “steerable” filters. The resulting method is both computationally efficient (with respect to the number of filter evaluations), robust to variation in pattern feature shape, and, unlike previous approaches, is applicable to a wide variety of pattern types. An application of the method is presented which uses a nearest-neighbour analysis to distinguish between uniform (defect-free) and non-uniform (strained, defect-containing) regions within imaged self-assembled domains, both with striped and hexagonal patterns.

Keywords: surface self-assembly, pattern recognition, shapelets, image processing, machine learning

1. Introduction

Modern microscopy techniques are producing an ever-increasing amount of high-resolution imaging data of self-assembled materials. There are thousands of images of such films in the self-assembly literature alone. One grand challenge in this area is to relate the structure and dynamics of materials as captured by imaging data to desired physical and chemical properties. To date, researchers have predominantly relied on purely qualitative techniques (*e.g.* visual inspection) or simple heuristic algorithms to interpret imaging data with this end goal in mind. However, such techniques i) cannot provide a quantitative description of the relationship between the imaging data and material properties and ii) do not scale to large amounts of data. Effectively using large amounts of imaging data to infer material properties requires *quantitative* characterization methods for the patterns in microscopy images (uniform regions, defects, *etc.*) that characterize the physical structure of the surface.

Recently developed methods for quantitative characterization, as shown in Figures 1a-1b, have yielded key fundamental insights into universality of self-assembly dynamics [1–4]. The methods use bond-orientational order theory [5], and have been applied mainly to studies of block copolymer (BCP) self-assembly on surfaces [1, 2, 6, 7]. They represent first steps toward solving the grand challenge. Figure 1b shows an example of such a method: given an image with both a known pattern (hexagonal) and convex pattern features, the method finds orientational relationships among these pattern features. These relationships can then be used in conjunction with bond-orientational order theory [5] to approximate local pattern orientation and identify defects. This type of quantification of surface order has been vital in the identification of pattern evolution mechanisms and defect kinetics [1–4, 6]. Furthermore, relationships resulting from this type of

*Corresponding author

Email addresses: dlizotte@uwaterloo.ca (Daniel Lizotte), nmabukhdeir@uwaterloo.ca (Nasser Mohieddin Abukhdeir)

analysis have since been shown to imply that pattern dynamics are universal, *i.e.* they are invariant with respect to chemical structure and physical interactions that drive pattern formation [3, 4]. Besides furthering fundamental understanding of self-assembly, these methods will be key enablers of the ultimate goal of controlling self-assembly to produce task-optimized material properties [8].

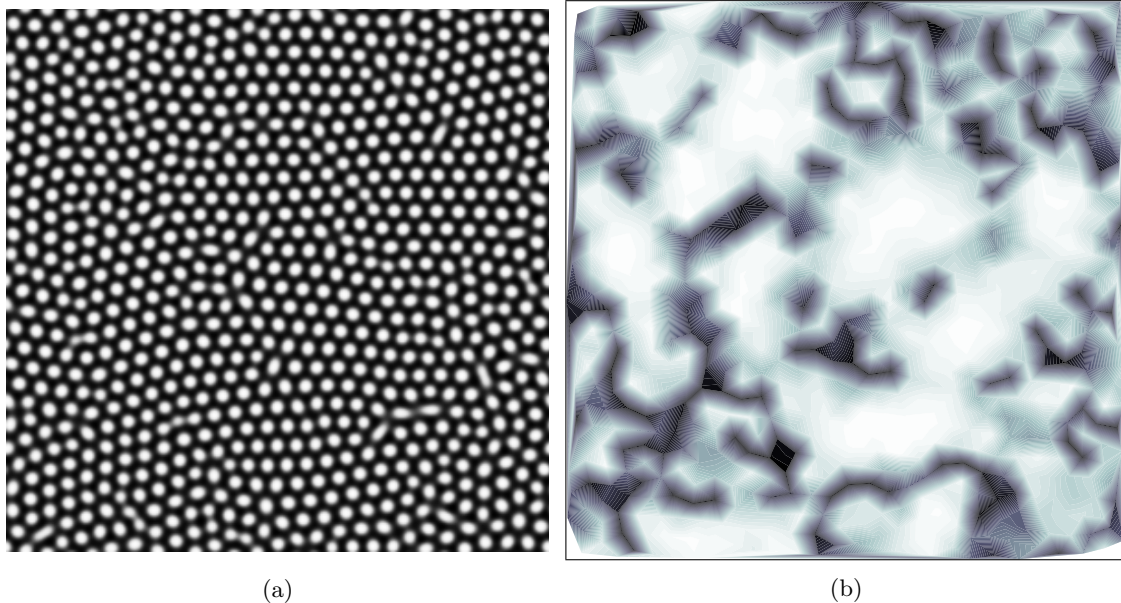


Figure 1: (a) Example of a (simulated) hexagonal self-assembled film from past work [9] where the field shows surface coverage of a species ranging from $0 \rightarrow 1$; (b) local hexagonal order resulting from applying the bond orientational order method to (a) which ranges from no order (black) to perfect order (white).

Despite its advantages, the bond-orientational order approach to self-assembly pattern analysis has several key limitations:

- *resolution* – bond-orientational order theory quantifies order at the pattern “feature” level, where pattern features are sub-domains which repeat in an ordered way. Typically, these features are larger than the resolution of the image, as is the case in Figure 1a, which results in a coarse resolution of local pattern order as shown in Figure 1b where interpolation is used.
- *convexity* – in order to compute *unique* nearest-neighbour “bonds” between pattern features, the features must be convex. This precludes the use of the method on stripe patterns and patterns in which features vary greatly in character (strained patterns).
- *uncertainty* – in order to compute nearest-neighbour “bonds” between pattern features, features must be uniquely identifiable. Typical experimental images of self-assembly phenomena involve nanoscale features which result in significant measurement uncertainty.

Furthermore, images frequently contain multiple regions that may or may not contain patterns, and that sometimes contain multiple patterns. Thus there is a clear need for robust, automated approaches to pattern recognition (“Is a pattern present in this image? Where within the image?”) and classification (“What type of pattern is present?”) for self-assembly imaging [10], in addition to a more detailed characterization (“How is the pattern oriented? Where are the defects?”) once these initial questions have been answered.

This work presents and demonstrates an analysis method for self-assembled surface imaging that is fundamentally different from past approaches for BCPs and that addresses the limitations described above. The basis of the method is a family of localized functions called *shapelets* [11]. Shapelets were originally developed to characterize images of galaxies ($\propto 10^{20}$ m) [11]; they are used here to characterize images of

nanopatterned surfaces ($\propto 10^{-9}$ m). It is demonstrated that, using simulation data of self-assembled surfaces, the presented approach is able to robustly determine *local* pattern characteristics, using an appropriate subset of shapelets [11, 12] and steerable filter theory [13], such as sub-domains that are well-ordered, strained, and/or have defects present.

2. Background

The analysis method synthesizes global pattern information derived from the discrete Fourier transform with local pattern information derived from projecting the image onto shapelet filters in a rotation-invariant way using steerable filter theory. Fourier analysis, shapelets, and steerable filter theory are reviewed below.

2.1. Fourier Analysis

The discrete Fourier transform (DFT) is a standard image analysis approach that can be used to quantify the presence of patterns or periodicity in an image. Given an image with intensity given by $f(x, y)$, the DFT of f is

$$\mathcal{F}\{f\}(u, v) = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x, y) \exp -i \left(\frac{2\pi ux}{X} + \frac{2\pi vy}{Y} \right) \quad (1)$$

which transforms the image from the spatial domain to the frequency domain. The resulting coefficients $\mathcal{F}\{f\}(u, v)$ of the Fourier modes over a discrete set of wave vectors indexed by u and v characterize both wavelength and orientation [14] of all periodic image patterns. While the DFT can be computed efficiently, the coefficients provide only *global* information in the sense that the Fourier modes span the whole spatial domain, *i.e.*, they are not localized in space. Thus, this decomposition can only determine the presence of domain-wide periodic components in the image and recover their characteristic wavelengths.

Figure 2a shows the resulting frequency domain representation of the image from Figure 1a, and Figure 2b shows the radially averaged spectral density. For images of simple uniformly-ordered domains (e.g. those with a single orientation) the DFT provides sufficient information about the type of pattern and its orientation to fully characterize the pattern. However, this simplicity is rarely observed in experimental imaging of self-assembled surfaces. In domains that are not well-ordered, the output of the DFT reveals only the presence of periodic structure within the image and characteristic length scales of that structure; it does not reveal *local* pattern structure. For example, the peaks in the radially averaged spectral density visible in Figure 2b reveal length scales of periodic patterns in Figure 1a, but neither they nor the full DFT in Figure 2a reveal the location of defects or grain boundaries.

2.2. Shapelets

Shapelets are a recently-proposed family of orthonormal basis functions which have been used for image analysis and are particularly suited to characterizing local pattern features. Shapelet analysis projects the image of interest onto basis functions of *fixed scale* and *varying shape*, shown in Figure 3. As with the windowed Fourier transform and wavelet decompositions [15], shapelet analysis [11] involves a linear decomposition of an image obtained by projecting it onto a set of localized orthonormal basis functions (Figure 3). The *polar shapelets* [12] are of particular interest as they possess rotational symmetries that are also present in images of self-assembled materials. They are given by,

$$B_{n,m}^{\odot}(r, \theta; \beta) = \beta^{-1} \chi_{n,m}(\beta^{-1}r) e^{-im\theta} \quad (2)$$

where n and m are nonnegative integers, and β is a characteristic length scale. The function $\chi_{n,m}$ is given by

$$\chi_{n,m}(r) = c_{n,m} r^m L_n^m(r^2) e^{-\frac{r^2}{2}} \quad (3)$$

where $c_{n,m}$ are constants, and $L_n^m(r) = \frac{r^{-m} e^r}{n!} \frac{d^n}{dr^n} (r^{m+n} e^{-r})$ are the associated Laguerre polynomials.¹ Examples for $n = 0$ and m from 0 through 6 are shown in Figure 3. Note that for $m > 0$ a polar

¹Massey et al. use alternative indices m', n' where $m' = m$ and $n' = 2n + m$. This requires additional constraints on m', n' .

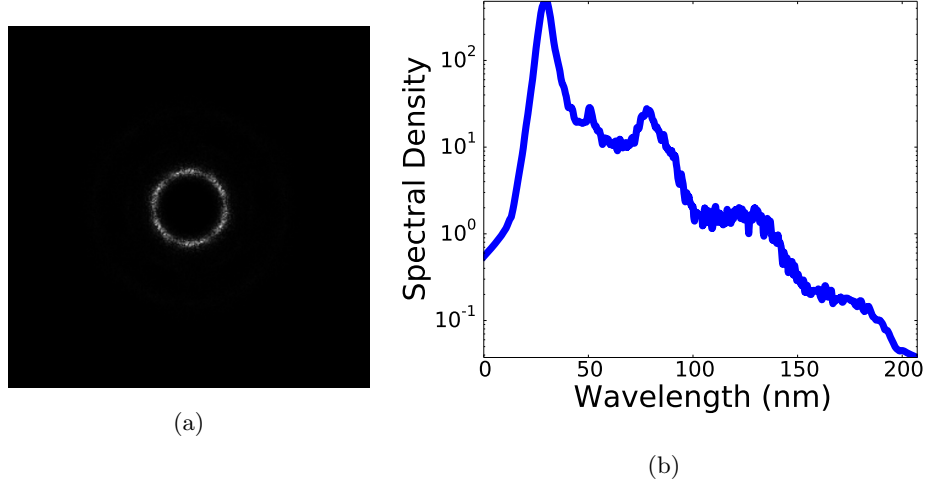


Figure 2: (a) Two-dimensional spectral density plot from the DFT of Figure 1a with the origin at the centre of the image; (b) radially averaged spectral density from (a).

shapelet has a real part and an imaginary part that is equal to a rotation of the real part by an angle of $-\pi/(2m)$. It is natural to define polar shapelets in polar coordinates; however, for convenience, define $B_{n,m}(x, y; \beta) \triangleq B_{n,m}^\odot(x^2 + y^2, \tan^{-1}(y, x); \beta)$ since images are expressed in Cartesian coordinates.

Standard shapelet analysis of an image is similar to that of other discrete transforms such as the Fourier and wavelet transforms. The image is decomposed into a linear combination of the basis functions,

$$f(x, y) = \sum_{n=0}^N \sum_{m=0}^M w_{n,m} B_{n,m}(x, y; \beta) \quad (4)$$

where each weighting coefficient $w_{n,m}$ is given by a discrete inner product or *correlation* of its corresponding shapelet $B_{n,m}(x, y; \beta)$ with the image function $f(x, y)$,

$$w_{n,m} = f \star B_{n,m} \triangleq \sum_{x'} \sum_{y'} f(x', y') B_{n,m}(x', y'; \beta). \quad (5)$$

The real (resp. imaginary) part of coefficient $w_{n,m}$ can be interpreted as a measure of similarity between the image and the real (resp. imaginary) part of $B_{n,m}$. The coefficient $w_{n,m}$ is termed the *response* for shapelet $B_{n,m}$. Because the shapelets are spatially localized functions, different translated versions of a shapelet have different responses. It is therefore natural to consider response as a function of image coordinates x and y

$$w_{n,m}(x, y) = \sum_{x'} \sum_{y'} f(x', y') B_{n,m}(x' - x, y' - y; \beta) \quad (6)$$

which gives the similarity between the image and a shapelet whose origin has been translated to image location (x, y) .

2.3. Steerable Filter Theory

Given a (typically real-valued) filter $F(x, y)$ expressed in Cartesian coordinates, it is often useful to know how that filter would respond if it (or the image) were rotated. Consider a version of the filter that has been rotated clockwise by a phase angle φ , $F(x, y; \varphi) = F(x \cos \varphi - y \sin \varphi, x \sin \varphi + y \cos \varphi)$. It is particularly useful to know the φ for which the filter response is maximal, as this angle contains information about pattern orientation: it gives the angle by which the pattern (or the filter) must be rotated in order to

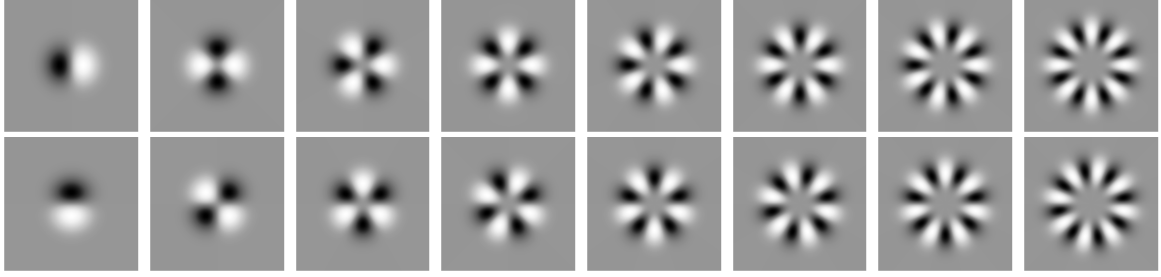


Figure 3: Surface plots of the real (top row) and imaginary (bottom row) components of polar shapelet functions (eqn. 2) with $n = 0$ and $m = 1 \rightarrow 8$. Image length-scale is normalized to 1 and $\beta = 0.55$.

achieve maximum similarity between the pattern and the filter. Freeman [13] shows if a filter is *steerable*, any rotated version of the filter can be expressed as a linear combination of a finite (and typically small) set of filters. This allows exact computation of the φ for which the filter response is maximal much more efficiently than brute-force techniques, which must explicitly compute a large number of filter orientations [16].

3. Results and Discussion

The successful shapelet-based analysis of self-assembly image data requires methods for determining an appropriate subset of shapelets, the optimal orientation of the shapelets, and the appropriate scales for the selected shapelets. Section 3.1 develops methods for each of these tasks, and Section 3.2 demonstrates the use of the developed methods on self-assembly data.

3.1. Sets of Steerable, Scale-Optimized Shapelets

In order to use shapelets for pattern analysis, appropriate shapelet sets and optimal scales are first determined based on “prototypical” uniform patterns that approximate real surface self-assembly imaging data, but that have a convenient parametric form. The uniform patterns are expressed in terms of a two-dimensional Fourier series [17],

$$\rho(\mathbf{x}) = \sum_{n=0}^N a_n \exp(i\mathbf{k}_n \cdot \mathbf{x}) \quad (7)$$

where the constants a_n are related to the magnitude of the pattern modulation and \mathbf{k}_n are the basis vectors of the pattern. For one-mode approximations of stripe and hexagonal patterns of interest, the basis vectors are [17],

$$\begin{aligned} \mathbf{k}_1 &= \frac{2\pi}{\lambda} \mathbf{e}_2 \\ \mathbf{k}_2 &= \frac{2\pi}{\lambda} \left(\frac{\sqrt{3}}{2} \mathbf{e}_1 - \frac{1}{2} \mathbf{e}_2 \right) \\ \mathbf{k}_3 &= \frac{2\pi}{\lambda} \left(\frac{-\sqrt{3}}{2} \mathbf{e}_1 - \frac{1}{2} \mathbf{e}_2 \right). \end{aligned} \quad (8)$$

where $\mathbf{e}_1 = (1, 0)^\top$ and $\mathbf{e}_2 = (0, 1)^\top$. For a stripe pattern, $a_1 \neq 0, a_2 = a_3 = 0$, and for a hexagonal pattern $a_1 = a_2 = a_3 \neq 0$. The quantity $2\pi/\lambda$ is the wavenumber for the pattern length scale (wavelength) λ , which corresponds to the peak shown in Figure 2a.

In order to select a minimal set of shapelets that respond strongly when applied to uniform stripe and hexagonal patterns, or in general any surface pattern, the sub-set of shapelets should have the following properties:

1. It should contain shapelets with the same fundamental rotational symmetries as the pattern of interest. Stripe patterns have subregions with 1- and 2-fold symmetry; hexagonal patterns have subregions with 1-, 2-, 3-, and 6-fold symmetry.
2. The response magnitude of the shapelets should be *invariant* with respect to rotations of the pattern.
3. The shapelets should respond most strongly to the dominant pattern frequencies.

3.1.1. Pattern Symmetries

Referring to Figure 3, a convenient property of shapelets of order (m, n) with $n = 0$ and $m > 0$ is that they have s -fold rotational symmetries for s corresponding to all numbers that divide m , plus the trivial 1-fold symmetry. Thus, at minimum, shapelets up to and including order $m = 2$ are necessary for analysis of stripe patterns, and shapelets up to order $m = 6$ are necessary for hexagonal patterns. For simplicity, shapelets up to order $m = 6$ are used in all example analyses, which form an overcomplete set of shapelets for stripe patterns and a minimal set for hexagonal patterns. The “redundant” information that this set provides in the striped pattern case did not prove to be problematic.

3.1.2. Rotational Invariance

In order to produce an analysis that is invariant to pattern rotations, for each shapelet in the set, the rotation of the shapelet that produces the largest response is determined. This is illustrated in Figure 4c. The brute-force, and thus computationally inefficient, approximate approach to finding the optimal rotation would be to determine the shapelet response for a large number of rotations of the shapelet and then select the rotation with the maximal response. This has two drawbacks: (i) the large number of shapelet responses must be evaluated at each pixel and (ii) the solution is not exact.

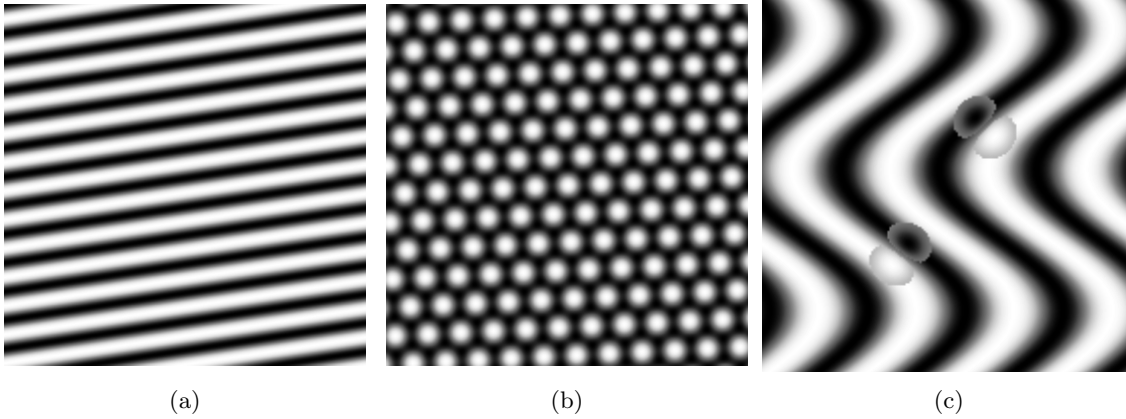


Figure 4: Plots of one-mode approximations of a (a) stripe and (b) hexagonal pattern using eqns. 7-8; (c) Schematic of rotations of a shapelet applied to a nonuniform stripe pattern.

Rather than solve for the optimal rotation approximately, an exact solution is derived from the fact that the shapelets are steerable.

Lemma 1. Let $B_{n,m}(x, y; \beta, \varphi) = B_{n,m}(x \cos \varphi - y \sin \varphi, x \sin \varphi + y \cos \varphi; \beta)$ be a shapelet as defined in (2) that has been rotated clockwise through a phase angle φ as described in Section 2.3. Then

$$\Re[B_{n,m}(x, y; \beta, \varphi)] = \cos(m\varphi) \Re[B_{n,m}(x, y; \beta)] + \sin(m\varphi) \Im[B_{n,m}(x, y; \beta)]. \quad (9)$$

Proof. (Sketch) Note from the polar form of shapelets that $B_{n,m}^\circ(r, \theta; \beta)$ is of the form $h(r)e^{-im\theta}$, and that a version rotated by an angle φ is of the form $h(r)e^{-im(\theta+\varphi)}$. The lemma follows from applying standard trigonometric identities to the complex exponential part of B° and simplifying. \square

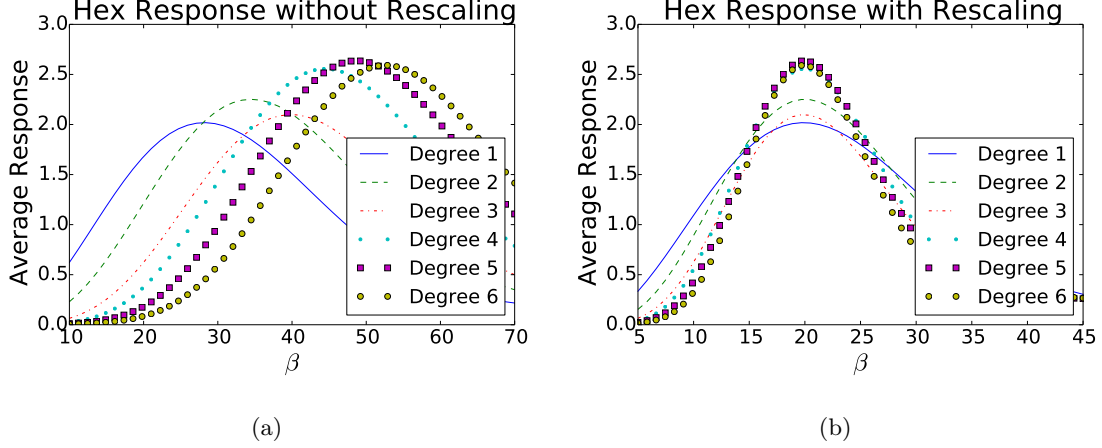


Figure 5: Shapelet filter response versus β (a) before and (b) after rescaling of β with respect to λ . Responses shown are correlations with the uniform pattern given by eqns. 7-8.

Steerable forms of each shapelet were formulated using Lemma 1 (eqn. 9), which are shown for $\varphi = 0$ in Figure 3 (highlighting of background figure). The steerable forms were then used to determine the optimal rotation angle for each shapelet at each image location.

Lemma 2. Let $w_{0,i} = f \star B_{0,i}(\cdot, \cdot; \beta)$, and define $\varphi_{0,i}^* = \arg \max_{\varphi} \Re[f \star B_{0,i}(\cdot, \cdot; \beta, \varphi)]$ and $w_{0,i}^* = \Re[f \star B_{0,i}(\cdot, \cdot; \beta, \varphi_{0,i}^*)]$. Then

$$\varphi_{0,i}^* = \arg w_{0,i}, \quad w_{0,i}^* = |w_{0,i}|. \quad (10)$$

Proof. (Sketch.) Since $\Re[B_{n,m}(x, y; \beta, \varphi)]$ is continuous in φ , it suffices to take the derivative, equate it to zero, and check second-order optimality conditions to find the optimal rotation and magnitude. Note that $\arg w_{0,i}$ is one of a countably infinite set of solutions to the optimal rotation problem. \square

Here, $\varphi_{0,i}^*$ is the shapelet orientation at which the real part of the *steered* shapelet response is maximal and $w_{0,i}^*$ is the value of the response at that orientation. In the above, the dependence of $\varphi_{0,i}^*$ and $w_{0,i}^*$ on x, y , and β is suppressed in the notation for clarity, but the lemma immediately applies to translated and scaled versions of shapelets as well. As desired, the rotation-optimized response $w_{0,i}^*$ is invariant to rotations of the pattern.

3.1.3. Scaling

The scale of a shapelet, given by the parameter β , also affects shapelet response. In order to ensure that the selected shapelets respond strongly to the pattern of interest and therefore do *not* respond strongly to pattern defects, their length scales are tuned to the dominant pattern frequency. Fixing a pattern, location (x, y) , and optimal orientation φ^* of a shapelet, the shapelet response is given by the correlation of the shapelet with the image function f ,

$$w_{n,m}^*(\beta) = f \star \Re[B_{n,m}(\cdot, \cdot; \beta, \varphi^*)] \triangleq \sum_{x'} \sum_{y'} f(x', y') \Re[B_{n,m}(x', y'; \beta, \varphi^*)], \quad (11)$$

which is only a function of β ; examples for different (n, m) are shown in Figure 5.

In order to ensure that the chosen shapelets respond optimally to the target pattern of interest, β is chosen so as to maximize the function $w_{C_1}^*$ (and $w_{C_2}^*$ and so on). For a given pattern length scale λ , the maximal shapelet response was found to *not* be at $\beta = \lambda$, but rather at $\beta = C\lambda$, with C depending on the order of the shapelet. Appropriate constants C were found for each shapelet using grid search on R ; these are given in Table 1. All analyses that follow use this rescaling.

Table 1: Coefficient values for $\beta = C\lambda$ for shapelets up to order 6.

| (\mathbf{m}, \mathbf{n}) | C | (\mathbf{m}, \mathbf{n}) | C |
|----------------------------|-------|----------------------------|-------|
| (1,0) | 1.418 | (2,0) | 1.725 |
| (3,0) | 2.003 | (4,0) | 2.224 |
| (5,0) | 2.439 | (6,0) | 2.640 |

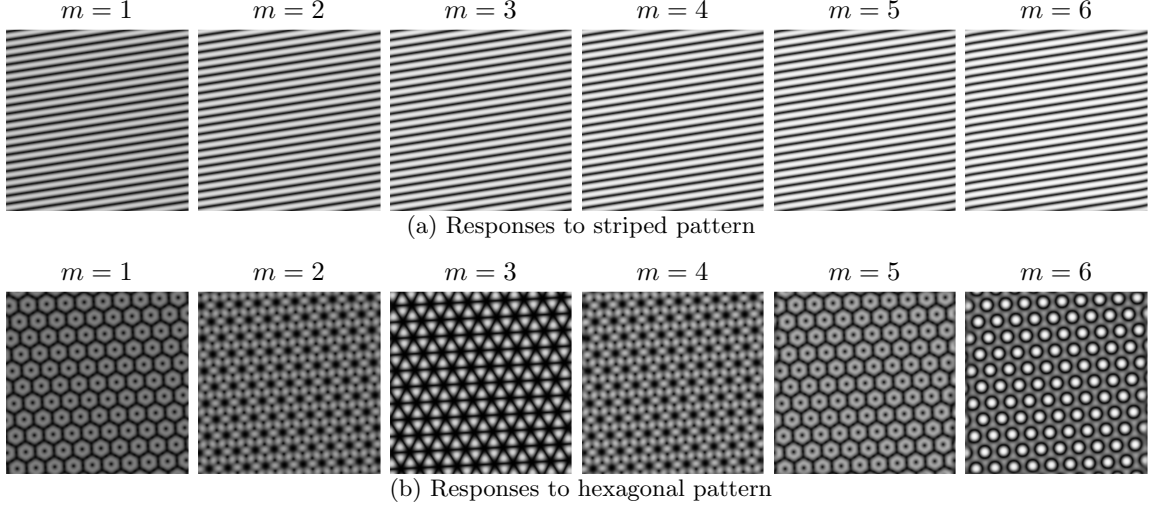


Figure 6: Responses of steerable polar shapelet filters applied to uniform one-mode approximations of a (a) stripe and (b) hexagonal pattern from eqns. 7-8.

Figure 6 shows rotation-optimized local pattern response values resulting from application of the scale-optimized steerable shapelets to uniform (i) stripe and (ii) hexagonal patterns using the uniform pattern given by eq. 7. The shapelet responses shown in Figure 6a and Figure 6b reveal the locations in the pattern that have different rotational symmetries. For example, the $m = 1$ shapelet applied to the striped pattern reveals where the pattern has only 1-fold symmetry: this occurs at the boundaries between stripes. At such locations, the pattern is only self-similar when rotated through a full 2π . The $m = 2$ shapelet, on the other hand, responds at peaks and troughs in the stripe pattern, where a rotation of π results in a self-similar pattern. As the shapelet order is increased beyond $m = 2$, almost no new local pattern information is extracted; the rotation-optimized responses of the shapelets with higher-order symmetry are very similar to the $m = 1$ or $m = 2$ case.

The $m = 1$ response applied to the hexagonal pattern responds at locations near the “edge” of a pattern mode, which have only 1-fold symmetry. The $m = 2$ and $m = 4$ shapelets respond strongly to areas that are midway along a line joining two pattern modes; these locations have 2-fold symmetry. The $m = 3$ shapelet responds very strongly to the 3-fold symmetry at “saddle points”—that is, at points in the pattern that are equidistant from three nearby modes, and finally the $m = 6$ shapelet responds strongly at pattern modes where there is 6-fold rotational symmetry. Note that the hexagonal pattern lacks any 5-fold symmetry; thus the $m = 5$ shapelet, which has only 5-fold and 1-fold symmetry responds only at pattern locations with 1-fold symmetry which are also identified by the $m = 1$ shapelet.

The next section explains how the responses of different shapelets can be combined into a useful quantitative analysis of the underlying image.

3.2. Application to Self-Assembly Imaging

Surface self-assembly imaging typically involves surfaces with patterns that are non-uniform and pattern features that are not well-approximated using a one-mode assumption. Figures 7 and 8 shows example images

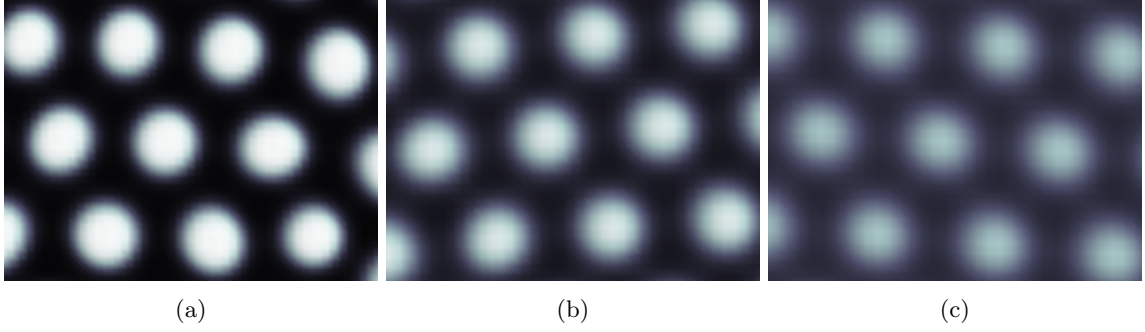


Figure 7: Examples of hexagonal surface self-assembly with features of varying character: (a) sharp interface, (b) semi-diffuse interface, (c) diffuse interface. Figures taken from ref. [9]

of two-dimensional surface self-assembly where non-uniform stripe and hexagonal patterns are present (taken from [9]). In Figure 7 the pattern features themselves vary in shape as is shown in Figure 7. In Figures 8a-8b multiple quasi-uniform subdomains, or “grains”, are present with defect regions (grain boundaries) at the interface between them. In order to test the presented shapelet-based method on these realistic patterns, a guided machine learning approach was used to classify regions with uniform patterns from those with defects.

The *response space* is defined as $\mathbf{r} \in \mathbb{R}^p$ where p is the number of steerable shapelet filters used to quantify the pattern. Thus at each point in the image (x, y) , a response vector \mathbf{r} is computed,

$$\mathbf{r}(x, y) = \frac{[w_{0,1}^*(x, y), w_{0,2}^*(x, y), \dots, w_{0,p}^*(x, y)]^T}{\|[w_{0,1}^*(x, y), w_{0,2}^*(x, y), \dots, w_{0,p}^*(x, y)]\|_2} \quad (12)$$

consisting of the shapelet responses $w_{0,j}^*(x, y)$ under the optimal orientation for image location (x, y) from eqn. 10. Given a user-specified set of coordinate pairs (i.e. pixel locations) \mathcal{R} of a defect-free subdomain of the image, at any location of interest (x', y') in the image the *response distance* may be defined from the pixel (x', y') to the reference set.

$$d_{\mathbf{r}}((x', y'), \mathcal{R}) = \min_{(x, y) \in \mathcal{R}} \|\mathbf{r}(x', y') - \mathbf{r}(x, y)\|_2, \quad (13)$$

where $d_{\mathbf{r}}((x', y'), \mathcal{R})$ is the Euclidean distance between the response vector at the location of interest and the closest response vector in the reference set.

The response distance encapsulates how different the image is at location (x', y') from the reference set in terms of the relative shapelet responses. It serves to highlight areas in the image where defects are present or where no pattern is present. Such areas have response vectors that have larger $d_{\mathbf{r}}((x', y'), \mathcal{R})$ from those where no defects are present.

This application of the steerable shapelets method was applied to the stripe and hexagonal self-assembled domain images shown in Figure 8. The characteristic pattern wavelength λ found through the maximum peak of the spectral density (inset of Figure 8) was used to determine the appropriate shapelet scale factors as described in Section 3.1. Figures 8c-8d illustrate the normalized Euclidean distance of the response vectors at each pixel with respect to the response vector of the uniform domain shown in Figure 8. In this figure, intensity is inversely proportional to $d_{\mathbf{r}}(\cdot, \mathcal{R})$ for the given quasi-uniform reference set, which clearly reveals the locations of defects in the image. Note that response distance is *invariant* to pattern rotations, because the elements of the response vectors are invariant to pattern rotations.

Pattern defects are of two general types: translational and orientational. These are referred to as dislocation and disclination defects [18], respectively, as shown in Figure 8. In stripe patterns, dislocations correspond to regions where a stripe feature begins (+) or ends (-). In hexagonal patterns, dislocations correspond to the beginning (+) or end (-) of a row of hexagonal features. Orientational defects, disclina-

tions, are manifested in a rapid transition from one pattern orientation to another. In stripe patterns, the majority of disclinations involve a $\frac{\pi}{2}$ rotation and, in hexagonal patterns, they involve a $\frac{\pi}{6}$ rotation.

The steerable shapelets method results, shown in Figure 8, show a direct relation between areas of strong response and quasi-uniform areas in the original pattern images. Areas where response is minimal corresponds to one of three localized cases: (i) the presence of defects, (ii) large strain of the pattern (stripe curvature and or dilation/compression), (iii) deviation of the pattern feature from the one-mode approximation.

With respect to defects present in both images, the image analysis results show good agreement with visual inspection of local topology in the original image. In areas of large strain of the pattern, which are typically also in the locality of defects, the shapelet response decays smoothly. This could be considered a drawback in that the method does not strongly distinguish between defect “core” regions and the region of strain surrounding the core. Alternatively, resolving the entirety of the region influenced by a single defect, or cluster of defects, likely has some significant in relating the pattern quality to material properties. Finally, in both images there are pattern features that strongly deviate from the one-mode approximation of the pattern. In the stripe pattern there are regions with convex circular shape and in the hexagonal pattern there are regions with lamellar-like features. The method is both robust in the presence of these features and strongly responds to their presence.

4. Conclusions

A method for quantitative analysis of surface self-assembly imaging was presented and applied to images of stripe and hexagonal ordered domains. A set of orthogonal functions, shapelets, were shown to be useful as filters which respond optimally to surface patterns with n -fold symmetry n is the order of the shapelet. Steerable formulations of the shapelet functions were derived using steerable filter theory and used to efficiently compute the filter rotation which yields maximal response. The utility of the steerable shapelet filter approach was demonstrated on uniform stripe and hexagonal patterns. Furthermore, realistic nonuniform surface patterns were analyzed using the presented steerable shapelet method through guided machine learning. This approach is able to *quantitatively* distinguish between uniform (defect-free) and non-uniform (strained, defects) regions within the imaged self-assembled domains. The presented method is both computationally efficient, requiring only two shapelet evaluations per steerable shapelet, and robust in the presence of variation in pattern feature shape. Finally, the shapelet-based method provides significantly enhanced resolution (pixel-level) compared to the bond-orientational order method (feature-level).

Acknowledgements

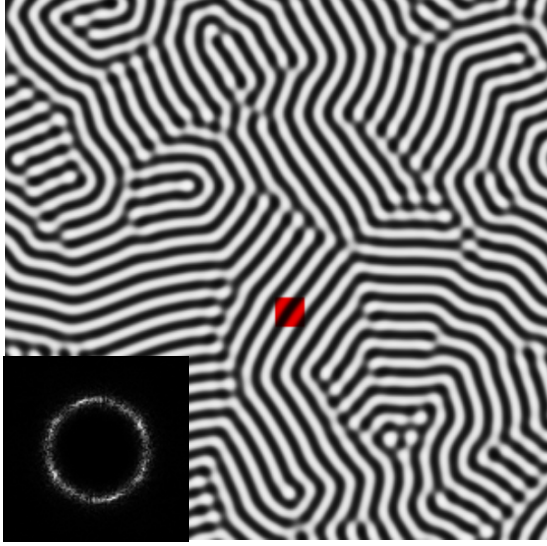
This work was made possible by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the facilities of the Shared Hierarchical Academic Research Computing Network (SHARC-NET).

References

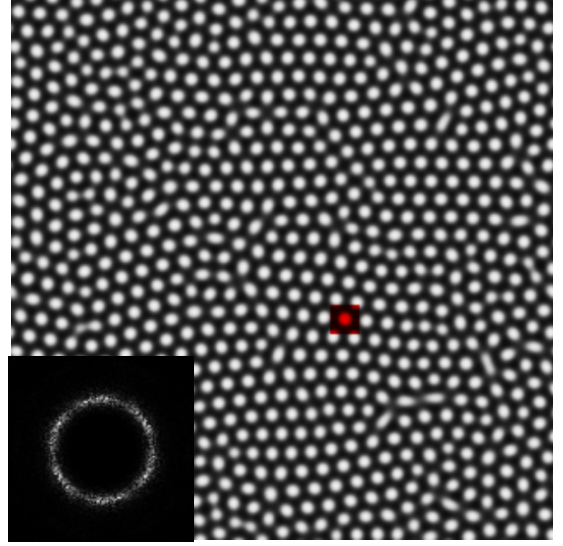
References

- [1] C. Harrison, D. H. Adamson, Z. Cheng, J. M. Sebastian, S. Sethuraman, D. A. Huse, R. A. Register, P. M. Chaikin, Mechanisms of Ordering in Striped Patterns, *Science* 290 (5496) (2000) 1558–1560. doi:10.1126/science.290.5496.1558.
- [2] C. Harrison, D. Angelescu, M. Trawick, Z. Cheng, D. Huse, P. Chaikin, D. Vega, J. Sebastian, R. Register, D. Adamson, Pattern Coarsening in a 2D Hexagonal System, *EPL (Europhysics Letters)* 67 (2004) 800.
- [3] N. M. Abukhdeir, A. D. Rey, Defect kinetics and dynamics of pattern coarsening in a two-dimensional smectic-A system, *New Journal of Physics* 10 (6) (2008) 063025 (17pp). URL <http://stacks.iop.org/1367-2630/10/063025>
- [4] N. M. Abukhdeir, D. G. Vlachos, Nanoscale Surface Pattern Evolution in Heteroepitaxial Bimetallic Films, *ACS Nano* 5 (9) (2011) 7168–7175. arXiv:<http://pubs.acs.org/doi/pdf/10.1021/nn201979q>, doi:10.1021/nn201979q. URL <http://pubs.acs.org/doi/abs/10.1021/nn201979q>
- [5] K. J. Strandburg (Ed.), *Bond-orientational Order in Condensed Matter Systems*, Springer-Verlag, 1992.

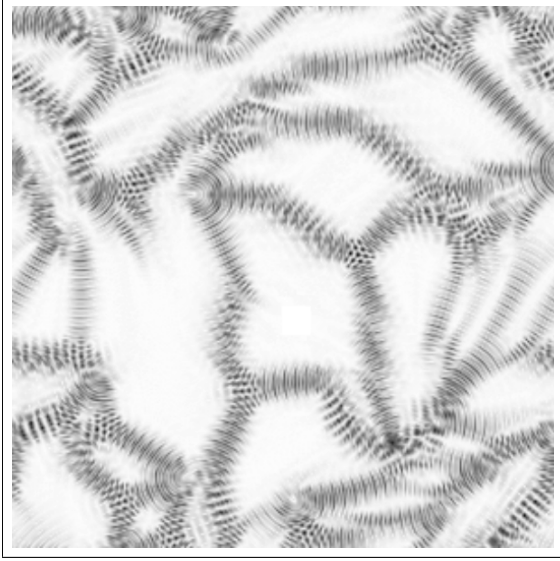
- [6] Y. Yokojima, Y. Shiwa, Hydrodynamic Interactions in Ordering Process of Two-dimensional Quenched Block Copolymers, *Phys. Rev. E* 65 (5) (2002) 056308. doi:10.1103/PhysRevE.65.056308.
- [7] D. A. Vega, C. K. Harrison, D. E. Angelescu, M. L. Trawick, D. A. Huse, P. M. Chaikin, R. A. Register, Ordering Mechanisms in Two-dimensional Sphere-forming Block Copolymers, *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* 71 (6) (2005) 061803. doi:10.1103/PhysRevE.71.061803.
URL <http://link.aps.org/abstract/PRE/v71/e061803>
- [8] J. McGill, N. Abukhdeir, B. Ogunnaike, D. Vlachos, Analysis and control of heteroepitaxial systems, *Journal of Process Control* Submitted, Special Issue on Control of Self-Assembly.
- [9] N. M. Abukhdeir, D. G. Vlachos, M. Katsoulakis, M. Plexousakis, Long-time Integration Methods for Mesoscopic Models of Pattern-forming Systems, *Journal of Computational Physics* 230 (14) (2011) 5704 – 5715. doi:DOI:10.1016/j.jcp.2011.03.052.
URL <http://www.sciencedirect.com/science/article/B6WHY-52HYRKV-2/2/a4c2c3cfee2f2f0fbc812d3873c3d66>
- [10] S. Rehse, K. Mecke, R. Magerle, Characterization of the dynamics of block copolymer microdomains with local morphological measures, *Phys. Rev. E* 77 (5) (2008) 051805. doi:10.1103/PhysRevE.77.051805.
- [11] A. Refregier, Shapelets i. a method for image analysis, *Monthly Notices of the Royal Astronomical Society* 338 (1) (2003) 35–47. doi:10.1046/j.1365-8711.2003.05901.x.
URL <http://dx.doi.org/10.1046/j.1365-8711.2003.05901.x>
- [12] R. Massey, A. Refregier, Polar shapelets, *Monthly Notices of the Royal Astronomical Society* 363 (1) (2005) 197–210. doi:10.1111/j.1365-2966.2005.09453.x.
URL <http://dx.doi.org/10.1111/j.1365-2966.2005.09453.x>
- [13] W. T. Freeman, E. H. Adelson, The design and use of steerable filters, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (1991) 891–906.
- [14] R. Szeliski, *Computer vision: algorithms and applications*, Springer, 2011.
- [15] J. Dealy, K. Wissbrun, *Melt Rheology and Its Role in Plastics Processing - Theory and Applications*, Springer, 1999.
- [16] Z. Wang, J. Li, Y. Guo, S. Tang, J. Wang, Unique visualization of multiply oriented lattice structures using a continuous wavelet transform, *Computer Physics Communications* 184 (11) (2013) 2489 – 2493. doi:http://dx.doi.org/10.1016/j.cpc.2013.06.018.
URL <http://www.sciencedirect.com/science/article/pii/S0010465513002129>
- [17] G. H. Gunaratne, Q. Ouyang, H. L. Swinney, Pattern formation in the presence of symmetries, *Physical Review E* 50 (4) (1994) 2802.
- [18] J. D. Brock, Bond-Orientational Order, in: Strandburg [5].



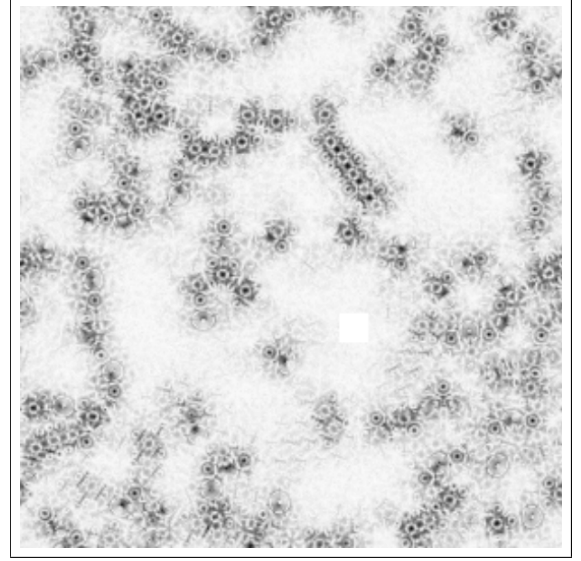
(a)



(b)



(c)



(d)

Figure 8: Examples of non-uniform (a) stripe and (b) hexagonal patterns from simulations of surface self-assembly (taken from ref. [9]) with inset spectral density plots; Results from applying the guided machine learning algorithm to the (c) stripe and (d) hexagonal patterns where the response distance (eqn. 13) was normalized to range from 0 (black) to 1 (white) and the user-specified set of pixels are highlighted in (a-b).